

<http://econometrics.blog.ir>

آموزش نرم افزارهای اقتصادسنجی

آموزش نرم افزارهای ایویوز، استاتا، لیزر، اکسل و میکروفیت



<http://econometrics.blog.ir>

آموزش نحوه انجام آزمون ناهمسانی واریانس داده های پانل دیتا در STATA

تهیه کننده: حسین خاندانی

کنترل و مشاوره: دکتر هادی امیری

آزمون ناهمسانی واریانس برای داده های پانل دیتا

در داده های تابلویی نیز مانند داده های سری زمانی می توان بحثهای مربوط به ناهمسانی واریانس بین جملات اختلال و همچنین خودهمبستگی را مطرح نمود.

*اصولا در صورتی که **دوره زمانی** مورد مطالعه در داده های پانل نسبت به **تعداد واحدهای انفرادی بیشتر** است، انتظار می رود بحث **خودهمبستگی** بین اجزای اخلاص موضوعیت داشته باشد.

در صورتی که **تعداد واحدهای انفرادی بیشتر باشد از **دوره زمانی** مورد مطالعه است، می توان انتظار داشت که اجزای اخلاص دارای **ناهمسانی واریانس** باشند. البته این مورد فقط یک قاعده سرانگشتی قبل از انجام آزمون است. (دکتر پرویز محمد زاده و همکاران، کاربرد استاتا در اقتصادسنجی)

✚ برای انجام آزمون ناهمسانی واریانس در استاتا مراحل زیر باید طی شود:

برای آزمون ناهمسانی واریانس بین جملات اخلاص دو مدل رگرسیون مقید (Restricted) و نامقید (UnRestricted) تخمین زده می شود.

در مدل مقید فرض همسانی واریانس یا فرض توزیع یکسان و مستقل جملات اختلال در نظر گرفته می شود، در حالیکه در مدل نامقید فرض بر یکسان نبودن واریانس جملات اخلاص بین واحدهای مقطعی (ناهمسانی واریانس) می باشد. در مرحله بعد با استفاده از دستور XTGLS در روش حداقل مربعات تعمیم یافته، هر دو مدل تخمین زده می شود و سپس بر اساس آماره آزمون نسبت راستنمایی (Likelihood Ratio) و با استفاده از فرمول محاسباتی زیر به آزمون فرضیه ناهمسانی- واریانس پرداخته می شود.

$$LR = 2(L_{UR} - L_R)$$



در رابطه فوق **LUR** لگاریتم راستنمایی مدل نامقید و **LR** لگاریتم راستنمایی در مدل مقید می باشد.

آزمون ناهمسانی واریانس در نرم افزار استاتا به صورت زیر انجام می شود:

UnRestricted Model:

`xtgls y x1 x2 x3, igls panels(heteroskedastic)`

پس از تخمین مدل غیر مقید نتایج را با دستور زیر ذخیره می نمایم.

`estimates store hetero`

در مرحله بعد **مدل مقید** بدون در نظر گرفتن ناهمسانی واریانس به صورت زیر تخمین زده می شود:

Restricted Model:

`xtgls y x1 x2 x3`

باز به مانند قبل، و این بار برای مدل مقید نتایج تخمین را با دستور زیر ذخیره می نمایم:

`estimates store hemo`

در مثال بالا **Y** متغیر وابسته و **X1, X2, X3** متغیرهای مستقل ما هستند.

در مرحله بعد برای تعیین محدودیتها یا درجه آزادی آماره آزمون، نسبت راستنمایی و درجه آزادی به صورت زیر تعریف می شود:

$$local\ df = e(N_g) - 1$$

در رابطه فوق $e(N_g)$ بیانگر تعداد گروهها یا تعداد واحدهای مقطعی می باشد.

در مرحله آخر: بعد از تعریف درجه آزادی نسبت راستنمایی، آماره آزمون LR به صورت زیر محاسبه می شود:

lrtest hetero hemo, df(df')

به علامت دو طرف df توجه کنید که به چه صورت است، خصوصا علامتی که با رنگ زرد مشخص شده است، چون نرم افزار در غیر این صورت خطا می‌زند. اگر نرم افزار از شما خطا گرفت روی خطا کلیک کنید تا نحوه صحیح آن را به شما اعلام کند (df')

با توجه به مدل و آزمون بالا می‌توان گفت که **مدل مقید (مدل دارای همسانی واریانس) در مدل **نامقید (دارای ناهمسانی واریانس)** آشیانه (Nested) کرده است.

پس از انجام آزمون ناهمسانی واریانس و **رد فرضیه صفر** در آماره آزمون نسبت راستنمایی لازم است **مدل نامقید** تخمین زده شود. زیرا مدل دارای ناهمسانی واریانس بوده است و برای تخمین مدل باید از روش **gls** با فرض در نظر گرفتن ناهمسانی واریانس استفاده کرد.

H_0 : همسانی واریانس

H_1 : ناهمسانی واریانس

بنابراین اگر **prob** آزمون **LR** و آماره **Chi2** زیر ۵ صدم شد فرضیه همسانی واریانس رد می‌شود و مدل رگرسیون دارای ناهمسانی واریانس است. در نتیجه باید از روش حداقل مربعات تعمیم یافته **xtgls** و مدل نامقید استفاده کرد.

در زیر نحوه انجام این آزمون در استاتا به صورت **تصویری** و مرحله به مرحله آورده شده است

قبل از هر کار ابتدا داده‌ها مربوطه را وارد نرم افزار نموده و **cross** و **years** را تعریف نموده و پس از تخمین مدل، مراحل زیر را در صورتی که مدل شما پانل است انجام دهید.



۱. مرحله اول دستور زیر را تایپ نمایید و اینتر را بزنید:

xtgls y x1 x2 x3, igls panels(heteroskedastic)

The screenshot shows the Stata/IC 12.0 [Results] window. The main window displays the following information:

```
.(6 variables, 110 observations pasted into data editor)
. xtset cross year
    panel variable: cross (strongly balanced)
    time variable: year, 2002 to 2012
    delta: 1 unit

. xtreg y x1 x2 x3

Random-effects GLS regression           Number of obs   =    110
Group variable: cross                  Number of groups =    10

R-sq:  within = 0.1822                  Obs per group:  min =    11
      between = 0.4515                    avg =           11.0
      overall  = 0.3746                    max =           11

corr(u_i, X) = 0 (assumed)              Wald chi2(3)    =   39.06
                                          Prob > chi2     =   0.0000
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
y					
x1	277.7626	264.7912	1.05	0.294	-241.2185 796.7438
x2	36499.87	6766.27	5.39	0.000	23238.22 49761.51
x3	-8187.032	4055.583	-2.02	0.044	-16135.83 -238.2361
_cons	-12196.53	7902.586	-1.54	0.123	-27685.31 3292.255
sigma_u	3912.8597				
sigma_e	6305.1931				
rho	.2780389	(fraction of variance due to u_i)			

Command: `xtgls y x1 x2 x3, igls panels(heteroskedastic)`

The right-hand side of the window shows the Variables list and Properties panel. The Variables list includes: year, cross, y, x1, x2, x3. The Properties panel shows details for variable x3: Name: x3, Label: (empty), Type: float, Format: %8.0g, Value Label: (empty), Notes: (empty). The Data panel shows: Filename: (empty), Label: (empty), Notes: (empty), Variables: 6, Observation: 110, Size: 2.04K.



۲. گام دوم با دستور زیر نتایج را ذخیر نمایند:

estimates store hetero

The screenshot shows the Stata/IC 12.0 interface. The main window displays the results of a cross-sectional time-series FGLS regression. The command window at the bottom shows the command `estimates store hetero` being executed. The Results window shows the following output:

```
Iteration 13: tolerance = .00001069
Iteration 14: tolerance = 4.983e-06
Iteration 15: tolerance = 2.257e-06
Iteration 16: tolerance = 1.002e-06
Iteration 17: tolerance = 4.375e-07
Iteration 18: tolerance = 1.888e-07
Iteration 19: tolerance = 8.068e-08
```

Cross-sectional time-series FGLS regression
Coefficients: generalized least squares
Panels: heteroakedaatic
Correlation: no autocorrelation

Estimated covariances	=	10	Number of obs	=	110
Estimated autocorrelations	=	0	Number of groups	=	10
Estimated coefficients	=	4	Time periods	=	11
Log likelihood	=	-1155.299	Wald chi2(3)	=	269.74
			Prob > chi2	=	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
x1	48.59394	133.3251	0.36	0.716	-212.7185 309.9064
x2	16483.49	5263.522	3.13	0.002	6167.178 26799.81
x3	16000.29	3040.524	5.26	0.000	10040.97 21959.61
_cons	-17352.43	4360.302	-3.98	0.000	-25898.47 -8806.395

Command window:
`estimates store hetero`



۳. گام سوم رگرسیون مقید را تخمین بزنید

xtgls y x1 x2 x3

Stata/IC 12.0 - [Results]

File Edit Data Graphics Statistics User Window Help

Review

	y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
x1		130.3134	159.8956	0.81	0.415	-183.0762 443.7029
x2		17716.47	5425.475	3.27	0.001	7082.739 28350.21
x3		10674.17	3323.976	3.21	0.001	4159.3 17189.05
_cons		-11204.7	5023.662	-2.23	0.026	-21050.9 -1358.503

. xtgls y x1 x2 x3

Cross-sectional time-series FGLS regression

Coefficients: generalized least squares
Panels: homoskedastic
Correlation: no autocorrelation

Estimated covariances = 1 Number of obs = 110
Estimated autocorrelations = 0 Number of groups = 10
Estimated coefficients = 4 Time periods = 11
Wald chi2(3) = 121.25
Log likelihood = -1163.091 Prob > chi2 = 0.0000

	y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
x1		130.3134	159.8956	0.81	0.415	-183.0762 443.7029
x2		17716.47	5425.475	3.27	0.001	7082.739 28350.21
x3		10674.17	3323.976	3.21	0.001	4159.3 17189.05
_cons		-11204.7	5023.662	-2.23	0.026	-21050.9 -1358.503

Command

xtgls y x1 x2 x3

Variables

Variable	Label
year	
cross	
y	
x1	
x2	
x3	
_est_hemo	esample() fr
_est_hetero	esample() fr

Properties

Variables

Name	x3
Type	float
Format	%8.0g
Value Label	
Notes	

Data

Filename	
Label	
Notes	
Variables	8
Observations	110
Size	2.26K

C:\Users\hossein\Documents

CAP NUM OVR

ENG 6:28 PM



۴. نتایج را با دستور زیر ذخیره نمایید:

.estimates store hemo

The screenshot shows the Stata/IC 12.0 interface. The Command window contains the following commands:

```
1 xtset cross year  
2 xtreg y x1 x2 x3  
3 xtgls y x1 x2 x3, igls ...  
4 estimates store hemo  
5 estimates store hemo  
6 estimates store hemo  
7 xtgls y x1 x2 x3
```

The Results window displays the following output:

```
              _cons      -17352.43   4360.302   -3.98   0.000   -25898.47   -8806.395  
-----  
. estimates store hetero  
. estimates store hemo  
. estimates store hetero  
. xtgls y x1 x2 x3  
Cross-sectional time-series FGLS regression  
Coefficients: generalized least squares  
Panels:      homoskedastic  
Correlation: no autocorrelation  
-----  
Estimated covariances = 1      Number of obs = 110  
Estimated autocorrelations = 0      Number of groups = 10  
Estimated coefficients = 4      Time periods = 11  
Wald chi2(3) = 121.25  
Log likelihood = -1163.091      Prob > chi2 = 0.0000  
-----  
              y      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]  
-----+-----  
x1          130.3134   159.8956     0.81   0.415   -189.0762   443.7029  
x2          17716.47   5425.475     3.27   0.001   7082.739   28350.21  
x3          10674.17   3323.976     3.21   0.001   4159.3    17189.05  
_cons      -11204.7      5023.662     -2.23   0.026   -21050.9   -1358.503  
-----
```

The Command window shows the command `estimates store hemo` has been executed.



۵. تعریف درجه آزادی

$$Local\ df = e(N_g) - 1$$

The screenshot shows the Stata/IC 12.0 interface. The main window displays the results of a cross-sectional time-series FGLS regression. The command window shows the following commands and output:

```
. xtglm y x1 x2 x3
Cross-sectional time-series FGLS regression
Coefficients: generalized least squares
Panels:      homoskedastic
Correlation: no autocorrelation

Estimated covariances = 1      Number of obs = 110
Estimated autocorrelations = 0    Number of groups = 10
Estimated coefficients = 4      Time periods = 11
Log likelihood = -1163.091      Wald chi2(3) = 121.25
                                Prob > chi2 = 0.0000
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
x1	130.3134	159.8956	0.81	0.415	-183.0762 443.7029
x2	17716.47	5425.475	3.27	0.001	7082.739 28350.21
x3	10674.17	3323.976	3.21	0.001	4159.3 17189.05
_cons	-11204.7	5023.662	-2.23	0.026	-21050.9 -1358.503

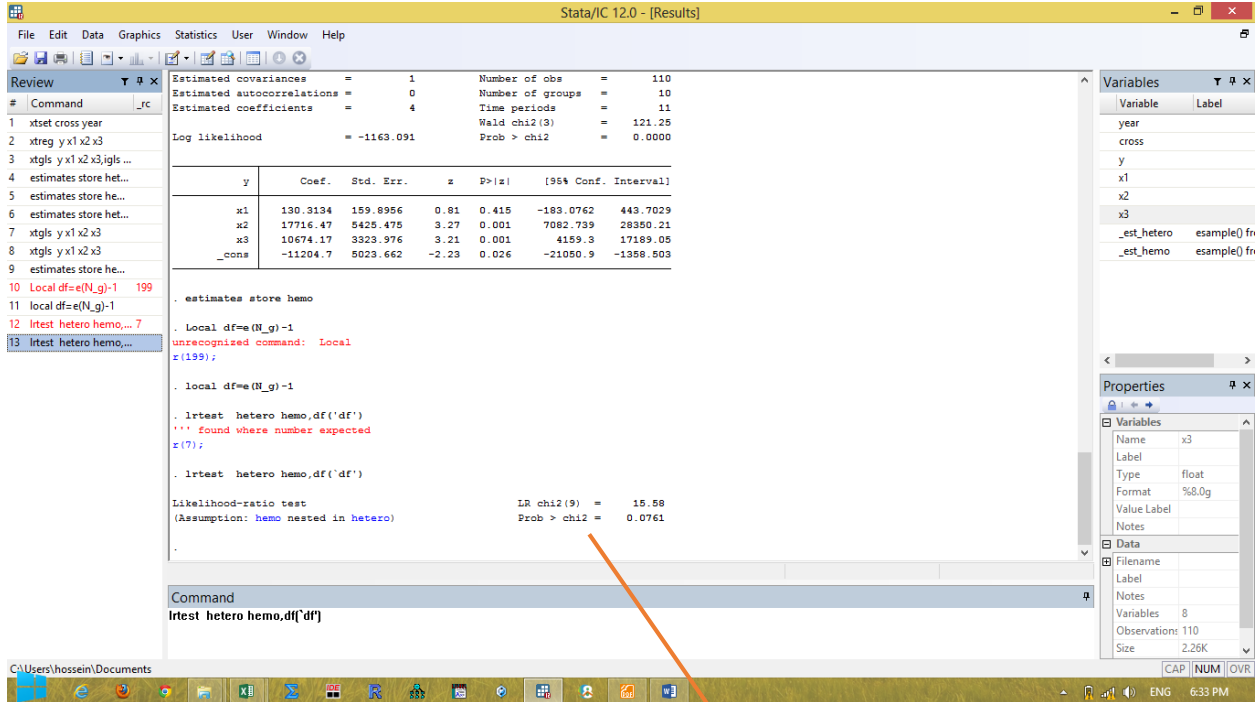
```
. estimates store hemo
. local df=e(N_g)-1
unrecognized command: Local
r(199);
. local df=e(N_g)-1
.
```

Command window:
local df=e(N_g)-1



۶. گام ششم وارد کردن دستور آزمون ناهمسانی واریانس $lrtest <$

lrtest hetero hemo,df(df')



```
. lrtest hetero hemo,df(df')

Likelihood-ratio test                      LR chi2(9) =    15.58
(Assumption: hemo nested in hetero)      Prob > chi2 =   0.0761

.

Command
```